Pergamon

Computers Math. Applic. Vol. 28, No. 8, pp. 87–98, 1994

0898-1221/94 $7.00 + 0.00

0898-1221(94)00172-3

Determination of Robot Location Using Generalized Cylindrical Object Shapes

JIANN-DER LEE*
Department of Electrical Engineering, Chang Gung College of Medicine and Technology
Tao-Yuan, Taiwan 333, R.O.C.
and
CHIN-HSING CHEN
Department of Electrical Engineering, National Cheng Kung University
Tainan, Taiwan 701, R.O.C.
and
JAU-YIEN LEE
Department of Electrical Engineering, Chang Gung College of Medicine and Technology
Tao-Yuan, Taiwan 333, R.O.C.

(Received March 1993; revised and accepted September 1993)

Abstract—In this paper, a new approach to the determination of robot location using the generalized cylindrical object, such as circular cylinders, elliptic cylinders, triangular cylinders, etc., are proposed. From a monocular image of the object, image processing and numerical analysis techniques are applied to extract the projection characteristics of the generalized cylindrical portion of the object, from which the position and the rotation parameters of a camera-mounted robot can be determined. Owing to the full linearity of the derivation, this approach can achieve high speed requirement. In addition, no prerequisite restriction is imposed on the image-grabbing process. Experimental results show that the location determination time is about 1.2 sec in a 33 MHz IBM compatible PC/AT computer system and the location error is less than 5% in average.

Keywords—Robot location determination, Generalized cylindrical shapes, Image processing, Vector analysis, Maximum-likelihood estimation.

NOMENCLATURE

C_1, C_2	two boundary curves of the top and bottom plane of a generalized cylindrical object	(X'_o, Y'_o, Z'_o)	the coordinate of the origin O in O'-$X'Y'Z'$ coordinate system
L_a, L_b	two normal lines perpendicular to the top and bottom planes of a generalized cylindrical object	R	the rotation matrix
h	the height of the generalized cylindrical object	f	the focal length of the camera
O-XYZ	the object coordinate system	λ_n	a ratio factor defined by $\overline{O'P_n}/\overline{O'P'_n}$, where P_n is a point in 3D space and P'_n is the corresponding point of it in the image plane
O'-$X'Y'Z'$	the camera coordinate system	$\hat{N}$	the direction of the object axis in O'-$X'Y'Z'$ coordinate system
U-V	the image coordinate system		
$(0, Y_c, Z_c)$	the coordinates of the camera lens center in O-XYZ coordinate system		

*Author to whom all correspondence should be sent.
This research is supported by the National Science Council, R.O.C. under Grant NSC81-0404-E006-570.

L_i	the distances between the camera lens center O' and P_i	S_1, S_2	two planes which are tangent to the generalized cylindrical object and pass through the camera lens center O'
K_i	a ratio factor		
ζ	a scaling constant	N_1, N_2	the normal vectors of S_1 and S_2

1. INTRODUCTION

For a robot to navigate automatically in various environments, it is important to determine its position with respect to some known objects in order to catch the objects or avoid collision. One approach to this so-called robot location problem is to simulate the human stereo vision. In this approach, binocular images are taken, and the image correspondence problem is solved to find the 3D range data.

Several remarkable approaches have been proposed to solve the problem alternatively by use of the "standard mark" [1–3] to simplify the problem and to reduce computation time. The major concept is to use special marks that include a wealth of geometric information under perspective projection such that robot location parameters can be easily computed from monocular images of the mark. In general, the methods using special marks can be categorized into two classes according to the condition that whether the camera optical axis goes through the center of the mark. When the camera optical axis is constrained to go through the center of the mark, the camera location can be simply represented by position parameters since the orientation is implicitly determined by the condition. In the other class, both the position and the orientation parameters are necessary. The specially designed marks include a diamond [1], a sphere with horizontal and vertical great circles [2], a tetragon [3], a triangle pattern consisting of two right-angled triangles of equal size [4], a house corner composed of three perpendicular lines [5], etc.

Curved objects are also encountered frequently in various environments. Haralick and Chu [6] solved camera parameters using the image of a parameterized curve such as a conic or a polygon. The solution procedure involves an iterating search to optimize the orientation parameters. The translation parameters are solved by an algebraic method based on the solutions of the orientation angles. Chen and Tsai [7] proposed a method for determining robot location by using surface patches of curve objects under the constraint that the focal length and the shape of the curved object are known in advance, and the equations of the curve can be described by a polynomial in two variables. However, since the focal length f (the actual distance between the camera lens center and the image plane) varies with the distance of the object to the camera, it is apparently unpractical to assume f to be known as a fixed value. Additionally, their method consumes much execution time by using iteration procedures.

In this paper, a new approach to robot location determination by using the generalized cylindrical object is proposed. Assumptions under the proposed approach include

(1) the desired object has a cylindrical portion, and
(2) the height of the cylindrical portion of the object is known.

Many objects found in indoor and outdoor environments satisfy the above constraints, such as cubes, cylinders, cans, furniture, machine parts, buildings, and so on. An example is shown in Figure 1. Each generalized cylindrical object used in this approach is composed of four space features—two planar curves (the curves C_1 and C_2 in Figure 1, for example), and two normal lines perpendicular to the top and bottom planes containing the curves C_1 and C_2 (L_a and L_b in Figure 1, for example). In contrast to the Tsai's method, the equations of the curves C_1 and C_2 can be any parameterized forms and need not be known in advance. In this approach, a monocular image of a generalized cylindrical object with known height is taken whenever the robot is to be located. The corresponding image features are then extracted and the side lines of the cylindrical shape are fitted in the least square error sensed by line equation. Next, two point-pairs are found from the upper and lower curves of the cylindrical shape by using the point-pair estimation strategy (described in Section 2) and robot location parameters can

then be determined uniquely by 3D vector analysis and simple algebraic computation. In the next section, the determination formulas of the robot location of the proposed approach are described. Image processing techniques required to obtain the desired features are presented in Section 3. Experimental results and discussions are included in Section 4, followed by conclusions and suggestions in Section 5.

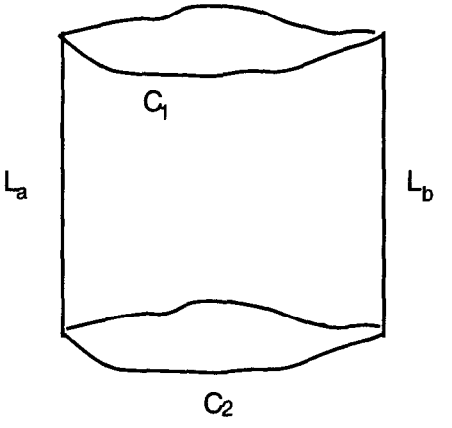

Figure 1. A generalized cylindrical object with the height known as h.

2. LOCATION DETERMINATION METHOD

A. The Mathematical Model of the System

The symbols used in this paper and their interpretations are shown in the Nomenclature. Figure 2 illustrates the imaging geometry of the system to be dealt with in this paper. In the figure, O-XYZ is the object coordinate system with the origin at the center of the curve formed by the lower edge of the generalized cylindrical object. The Z-axis coincides with the object axis and the X-Y plane is chosen to be located on the bottom surface of the object (see Figure 2). O'-$X'Y'Z'$ is the camera coordinate system with the origin at the camera lens center and the Y'-axis coinciding with the optical axis. Let the X'-axis and Z'-axis be parallel to the u-axis and v-axis, respectively. So, an image point located at (u, v) has the coordinates (u, f, v) in the camera coordinate system, where f is the focal length of the camera. Without loss of the generality, for any point P_n in the 3D space with the object coordinates (x, y, z), the camera coordinates (x', y', z') and the image coordinate (u, v) of its corresponding point P'_n in the image plane, the following equations have to be satisfied.

$$(x, y, z)^t = R \cdot (x', y', z')^t + (0, Y_c, Z_c)^t; \tag{1}$$

$$(x', y', z') = \lambda(u, f, v);$$

and

$$R = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix}, \qquad \text{(i.e., the rotation matrix).}$$

where λ is a ratio factor, $(0, Y_c, Z_c)$ are the coordinates of the camera lens center (robot location) in the object coordinate system. (From the viewpoint of the observer, it is reasonable to set the first coordinate of the translation vector to be zero because we can always find an object coordinates system with the property that the directions of the camera optical axis lies on the plane of Y-Z in the object coordinate system.)

To obtain an unique solution to the 3D location problem, it is well known that at least four points are needed. Therefore, in our approach, two point-pairs are extracted from the image of

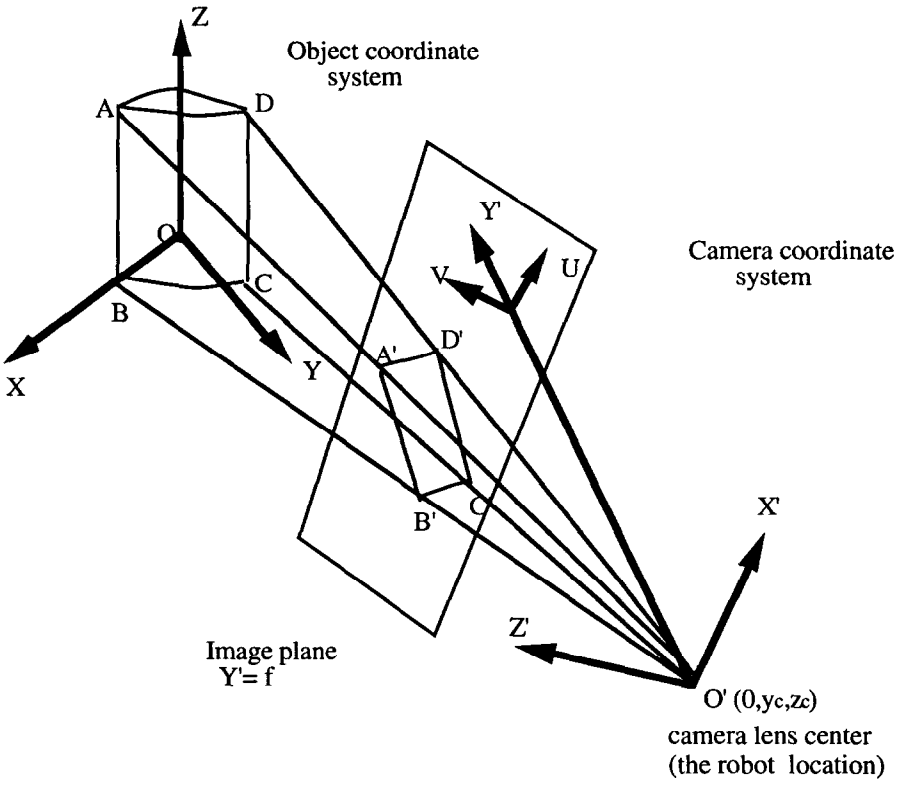

Figure 2. The camera viewing model of a generalized cylindrical object.

the cylindrical object. The so-called point-pairs are formed by two points, one at the upper curve and the other at the lower curve of the generalized cylindrical object, with their line segment parallel to the object axis. Let $[P_1, P_2]$ and $[P_3, P_4]$ denote two point-pairs. Since both the upper and the lower curves of the generalized cylindrical object are perpendicular to the object axis, the four line segments formed by the two point-pairs (i.e., $\overline{P_1P_2}$, $\overline{P_3P_4}$, $\overline{P_1P_3}$, and $\overline{P_2P_4}$) constitute a rectangle with P_1, P_2, P_3 and P_4 as its four vertices, and the length of $\overline{P_1P_2}$ (or $\overline{P_3P_4}$) is equal to the vertical height of the generalized cylindrical object (i.e., h). Note that P_1 and P_2 (or P_3 and P_4) have the same x and y coordinates and a difference h in z coordinate in the object coordinate system.

B. Point-Pair Estimation Strategy

Let P_1' be an image point on the upper curve, we now show how to determine P_2' such that P_1 and P_2 form a point-pair. Having known that P_2' must be located on the lower curve (i.e., the projection of the lower curve of the cylindrical object), we need only to test each point on the lower curve if it satisfies the definition. Assume that P_1' and P_2' have the image coordinates (u_1, v_1) and (u_2, v_2), respectively. According to the definition, $\overrightarrow{P_1P_2}$ must be parallel to the object axis, (i.e., $\overrightarrow{P_1P_2}//\hat{N}$), where $\hat{N}$ is the direction of the object axis in the camera coordinate (the derivation of $\hat{N}$ is described in the Appendix); we have

$$\lambda_2(u_2, f, v_2) - \lambda_1(u_1, f, v_1)//(a_3, b_3 f, c_3), \tag{2a}$$

where λ_n is a ratio factor defined by $\overline{O'P_n}/\overline{O'P_n'}$, $n = 1, 2$. Let $\lambda_1 = k\lambda_2$, where k is a variable. Replacing λ_1 in equation (2a) by $k\lambda_2$, we can get

$$\frac{\lambda_2(ku_1 - u_2)}{a_3} = \frac{\lambda_2 f(k-1)}{b_3 f} = \frac{\lambda_2(kv_1 - v_2)}{c_3}. \tag{2b}$$

Simplifying equation (2b), λ_2 can be deleted and the equations can be described by

$$a_3(k-1) = b_3(ku_1 - u_2), \tag{3a}$$

$$c_3(k-1) = b_3(kv_1 - v_2), \tag{3b}$$

$$c_3(ku_1 - u_2) = a_3(kv_1 - v_2). \tag{3c}$$

Our purpose is to determine (u_2, v_2) for prechosen (u_1, v_1) and given $\hat{N}$. From eqs. (3a)–(3c), we can obtain the following three expressions of k, respectively:

$$k = \frac{b_3u_2 - a_3}{b_3u_1 - a_3}, \tag{4a}$$

$$k = \frac{b_3v_2 - c_3}{b_3v_1 - c_3}, \tag{4b}$$

$$k = \frac{a_3v_2 - c_3u_2}{a_3v_1 - c_3u_1}. \tag{4c}$$

Note that it is impossible that $\hat{N}$ (the direction of the object axis) is a zero vector (i.e., it is impossible that $a_3 = b_3 = c_3 = 0$). To prevent k from being infinite, we choose (u_1, v_1) so that none of the denominators of equations (4a)–(4c) is equal to zero. If two of the three components of $\hat{N}$ are equal to zero (this may occur), then one of the denominators of equations (4a)–(4c) is equal to zero. (u_1, v_1) is chosen so that none of the other two denominators is equal to zero. Point (u_2, v_2) on the lower curve must satisfy equations (4a)–(4c). However, it is difficult to obtain (u_2, v_2) exactly due to the finite resolution of the camera. Therefore, we use the maximum-likelihood estimation [8] to determine (u_2, v_2). This maximum-likelihood estimation method is summarized as below: First, we define the discriminant Δ according to the following three cases:

Case 1. $|a_3| = \max\{|a_3|, |b_3|, |c_3|\}$. The discriminant is defined by

$$\Delta = \left| \frac{(b_3u_2 - a_3)/(b_3u_1 - a_3)}{(a_3v_2 - c_3u_2)/(a_3v_1 - c_3u_1)} - 1 \right|.$$

Case 2. $|b_3| = \max\{|a_3|, |b_3|, |c_3|\}$. The discriminant is defined by

$$\Delta = \left| \frac{(b_3u_2 - a_3)/(b_3u_1 - a_3)}{(b_3v_2 - c_3)/(b_3v_1 - c_3)} - 1 \right|.$$

Case 3. $|c_3| = \max\{|a_3|, |b_3|, |c_3|\}$. The discriminant is defined by

$$\Delta = \left| \frac{(b_3v_2 - c_3)/(b_3v_1 - c_3)}{(a_3v_2 - c_3u_2)/(a_3v_1 - c_3u_1)} - 1 \right|.$$

Second, compute the value Δ for each point on lower curve and the point with minimizes Δ is regarded as the optimal solution for (u_2, v_2).

By the same procedure described above, we can obtain another point-pair, say, $[P_3, P_4]$ as shown in Figure 2. Assume that the image coordinates of P_i', $i = 1 \ldots 4$, are (u_i, v_i), $i = 1 \ldots 4$, respectively. Thus P_i, $i = 1 \ldots 4$, have the camera coordinates $\lambda_i(u_i, f, v_i)$, $i = 1 \ldots 4$, respectively. Again, assume that P_i, $i = 1 \ldots 4$, have the object coordinates (x_1, y_1, h), (x_1, y_1, o), (x_2, y_2, h) and (x_2, y_2, o), respectively, where x_1, y_1, x_2 and y_2 are constants depending on the positions of P_1 and P_3. Since $\overrightarrow{P_1P_2} = \overrightarrow{P_3P_4}$, we have

$$\lambda_2(u_2, f, v_2) - \lambda_1(u_1, f, v_1) = \lambda_4(u_4, f, v_4) - \lambda_3(u_3, f, v_3). \tag{5}$$

From equation (5), we get

$$-u_1\lambda_1 + u_2\lambda_2 + u_3\lambda_3 = u_4\lambda_4, \tag{6a}$$

$$-\lambda_1 + \lambda_2 + \lambda_3 = \lambda_4, \tag{6b}$$

$$-v_1\lambda_1 + v_2\lambda_2 + v_3\lambda_3 = v_4\lambda_4. \tag{6c}$$

From equations (6a)–(6c), we can derive the relation between λ_1, λ_2, λ_3, and λ_4 as

$$\lambda_i = K_i \lambda_4, \qquad i = 1 \ldots 3, \tag{7}$$

where $K_i = \sigma_i/\sigma$, $i = 1 \ldots 3$, with

$$\sigma = \begin{vmatrix} -u_1 & u_2 & u_3 \\ -1 & 1 & 1 \\ -v_1 & v_2 & v_3 \end{vmatrix}, \qquad \sigma_1 = \begin{vmatrix} u_4 & u_2 & u_3 \\ 1 & 1 & 1 \\ v_4 & v_2 & v_3 \end{vmatrix},$$

$$\sigma_2 = \begin{vmatrix} -u_1 & u_4 & u_3 \\ -1 & 1 & 1 \\ -v_1 & v_4 & v_3 \end{vmatrix}, \qquad \sigma_3 = \begin{vmatrix} -u_1 & u_2 & u_4 \\ -1 & 1 & 1 \\ -v_1 & v_2 & v_4 \end{vmatrix}.$$

Since $(-u_1, -1, -v_1)$, $(u_2, 1, v_2)$ and $(u_3, 1, v_3)$ are noncoplanar, $\sigma \neq 0$ and K_1, K_2 and K_3 are unique. From equation (7), λ_1, λ_2 and λ_3 can be solved after λ_4 is solved.

Before solving λ_4, we have to solve the focal length f. Since $\overrightarrow{P_1P_3} \cdot \overrightarrow{P_1P_2} = 0$, we have

$$[\lambda_3(u_3, f, v_3) - \lambda_1(u_1, f, v_1)] \cdot [\lambda_2(u_2, f, v_2) - \lambda_1(u_1, f, v_1)] = 0. \tag{8}$$

Substituting equation (7) into equation (8), we get

$$[K_3u_3 - K_1u_1,\ (K_3 - K_1)\, f,\ K_3v_3 - K_1v_1] \cdot [K_2u_2 - K_1u_1,\ (K_2 - K_1)\, f,\ K_2v_2 - K_1v_1] = 0.$$

Thus,

$$f = \left[-\frac{(K_3u_3 - K_1u_1)(K_2u_2 - K_1u_1) + (K_3v_3 - K_1v_1)(K_2v_2 - K_1v_1)}{(K_3 - K_1)(K_2 - K_1)} \right]^{1/2}.$$

From the above equation, we learn that the focal length can be calculated with the image coordinate of two point-pairs chosen by point-pair estimation strategy, whenever a new image is grabbed. Additionally, no object coordinate value are needed in deriving the focal length because the information provided by two point-pairs is sufficient. (They can construct a rectangle-shaped standard mark and it is proved [8] that a unique solution always exists under this condition.)

Since $\overline{P_3P_4} = h$, we have

$$\|\lambda_3(u_3, f, v_3) - \lambda_4(u_4, f, v_4)\| = h, \tag{9}$$

where $\|\cdot\|$ denotes the Euclidean norm of a vector. Substituting equation (7) into equation (9), we get

$$\lambda_4 = \frac{h}{[(k_3u_3 - u_4)^2 + (k_3f - f)^2 + (k_3v_3 - v_4)^2]^{1/2}}.$$

Having found λ_4, the other three ratio factors λ_1, λ_2 and λ_3 can also be found from equation (7).

Let L_i, $i = 1 \ldots 4$ denote the distances between the camera lens center O' and P_i, $i = 1 \ldots 4$, respectively. That is

$$L_i = \overline{O'P_i} = \|\lambda_i(u_i, f, v_i)\|. \tag{10}$$

Therefore, L_1, L_2, L_3 and L_4 can be calculated from the above equations. The object coordinates of O' are $(0, Y_c, Z_c)$, and L_1 and L_2 are also given by

$$L_1 = \|(0, Y_c, Z_c) - (x_1, y_1, h)\|, \tag{11a}$$

$$L_2 = \|(0, Y_c, Z_c) - (x_1, y_1, 0)\|. \tag{11b}$$

Squaring equations (11a) and (11b), and subtracting one from the other, we can get

$$Z_c = \frac{L_2^2 - L_1^2 + h^2}{2h}.$$

From equation (1), the coordinate correspondences of P_1, P_2 and P_3 in the camera and object coordinate system are

$$\begin{aligned} P_1 &: \lambda_1(u_1, f, v_1) \longrightarrow (x_1, y_1, h), \\ P_2 &: \lambda_2(u_2, f, v_2) \longrightarrow (x_1, y_1, 0), \\ P_3 &: \lambda_3(u_3, f, v_3) \longrightarrow (x_2, y_2, h). \end{aligned}$$

Therefore,

$$\begin{bmatrix} x_1 & x_1 & x_2 \\ y_1 & y_1 & y_2 \\ h & 0 & h \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 u_1 & \lambda_2 u_2 & \lambda_3 u_3 \\ \lambda_1 f & \lambda_2 f & \lambda_3 f \\ \lambda_1 v_1 & \lambda_2 v_2 & \lambda_3 v_3 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ Y_c & Y_c & Y_c \\ Z_c & Z_c & Z_c \end{bmatrix}.$$

Rearranging the above equation for R_{31}, R_{32} and R_{33}, we have

$$\begin{bmatrix} \lambda_1 u_1 & \lambda_1 f & \lambda_1 v_1 \\ \lambda_2 u_2 & \lambda_2 f & \lambda_2 v_2 \\ \lambda_3 u_3 & \lambda_3 f & \lambda_3 v_3 \end{bmatrix} \cdot \begin{bmatrix} R_{31} \\ R_{32} \\ R_{33} \end{bmatrix} = \begin{bmatrix} h - Z_c \\ -Z_c \\ h - Z_c \end{bmatrix}. \tag{12}$$

Since $(\lambda_1 u_1, \lambda_1 f, \lambda_1 v_1)$, $(\lambda_2 u_1, \lambda_2 f, \lambda_2 v_2)$ and $(\lambda_3 u_3, \lambda_3 f, \lambda_3 v_3)$ are noncoplanar, we get

$$\begin{bmatrix} R_{31} \\ R_{32} \\ R_{33} \end{bmatrix} = \begin{bmatrix} \lambda_1 u_1 & \lambda_1 f & \lambda_1 v_1 \\ \lambda_2 u_2 & \lambda_2 f & \lambda_2 v_2 \\ \lambda_3 u_3 & \lambda_3 f & \lambda_3 v_3 \end{bmatrix}^{-1} \cdot \begin{bmatrix} h - Z_c \\ -Z_c \\ h - Z_c \end{bmatrix}.$$

To derive (X_o', Y_o', Z_o'), an arbitrary point (except P_2' and P_4') on the lower curve, say, P_5' with the image coordinates (u_5, v_5), is required as shown in Figure 3. The camera coordinates and object coordinates of P_5 are $\lambda_5(u_5, f, v_5)$ and $(x_5, y_5, 0)$, respectively, where λ_5 is the ratio factor defined by $\overline{O'P_5}/\overline{O'P_5'}$, and x_5 and y_5 depend on the location of P_5'. Here, λ_5 is an unknown to be solved. According to equation (1), we have

$$\begin{bmatrix} x_5 \\ y_5 \\ 0 \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \cdot \begin{bmatrix} \lambda_5 u_5 \\ \lambda_5 f \\ \lambda_5 v_5 \end{bmatrix} + \begin{bmatrix} 0 \\ Y_c \\ Z_c \end{bmatrix}.$$

Thus, we get

$$\lambda_5 u_5 R_{31} + \lambda_5 f R_{32} + \lambda_5 v_5 R_{33} + Z_c = 0,$$

and

$$\lambda_5 = -\frac{Z_c}{u_5 R_{31} + f R_{32} + v_5 R_{33}}.$$

Since $\overrightarrow{P_2O} \times \overrightarrow{P_2P_4} // \overrightarrow{P_2P_1}$, where $\times$ represents the outer product of two vectors, we have

$$\begin{vmatrix} i & j & k \\ X_o' - \lambda_2 u_2 & Y_o' - \lambda_2 f & Z_o' - \lambda_2 v_2 \\ \lambda_4 u_4 - \lambda_2 u_2 & \lambda_4 f - \lambda_2 f & \lambda_4 v_4 - \lambda_2 v_2 \end{vmatrix} = \zeta(\lambda_1 u_1 - \lambda_2 u_2, \lambda_1 f - \lambda_2 f, \lambda_1 v_1 - \lambda_2 v_2), \tag{13}$$

where ζ is a scaling constant. Since the determinant of the matrix formed by the three coplanar vectors $\overrightarrow{P_5O}$, $\overrightarrow{P_5P_2}$ and $\overrightarrow{P_5P_4}$ is equal to zero, we have

$$\begin{vmatrix} X_o' - \lambda_5 u_5 & Y_o' - \lambda_5 f & Z_o' - \lambda_5 v_5 \\ \lambda_2 u_2 - \lambda_5 u_5 & \lambda_2 f - \lambda_5 f & \lambda_2 v_2 - \lambda_5 v_5 \\ \lambda_4 u_4 - \lambda_5 u_5 & \lambda_4 f - \lambda_5 f & \lambda_4 v_4 - \lambda_5 v_5 \end{vmatrix} = 0. \tag{14}$$

From equations (13) and (14), we obtain

$$a_{11}X'_o + a_{12}Y'_o + a_{13}Z'_o = a_{14}, \tag{15a}$$
$$a_{21}X'_o + a_{22}Y'_o + a_{23}Z'_o = a_{24}, \tag{15b}$$
$$a_{31}X'_o + a_{32}Y'_o + a_{33}Z'_o = a_{34}. \tag{15c}$$

Next, using equations (15a)–(15c), (X'_o, Y'_o, Z'_o) can be obtained as below

$$X'_o = \begin{vmatrix} a_{14} & a_{12} & a_{13} \\ a_{24} & a_{22} & a_{23} \\ a_{34} & a_{32} & a_{33} \end{vmatrix} / \Omega, \qquad Y'_o = \begin{vmatrix} a_{11} & a_{14} & a_{13} \\ a_{21} & a_{24} & a_{23} \\ a_{31} & a_{34} & a_{33} \end{vmatrix} / \Omega,$$

$$Z'_o = \begin{vmatrix} a_{11} & a_{12} & a_{14} \\ a_{21} & a_{22} & a_{24} \\ a_{31} & a_{32} & a_{34} \end{vmatrix} / \Omega, \qquad \Omega = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

After (X'_o, Y'_o, Z'_o) is obtained, Y_c can be derived as below

$$Y_c^2 = \|\overline{OO'}\|^2 - Z_c^2,$$

where

$$\|\overline{OO'}\|^2 = X_o^{2'} + Y_o^{2'} + Z_o^{2'}.$$

From the above derivation, we have obtained the robot location $(0, Y_c, Z_c)$ relative to the generalized cylindrical object.

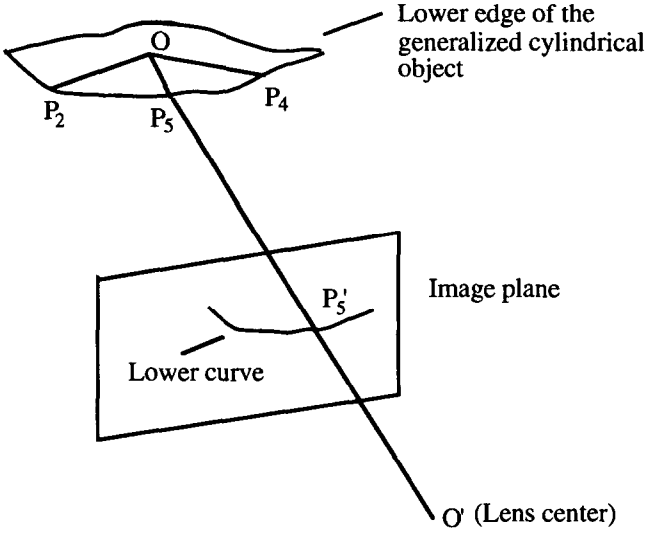

Figure 3. An additional point P'_5 with its 3D corresponding point P_5 located on the same curve with P_2 and P_4.

3. IMAGE PROCESSING TECHNIQUES

The required features for robot location determination are the equations of lines L_a and L_b and the points on the upper and the lower curves, as shown in Figure 1. Five cylindrical objects have been made for testing, including a triangular cylinder, a rectangular cylinder, an elliptic cylinder, a circular cylinder and a hexangular cylinder. To extract the required features (i.e., two point-pairs), the image processing techniques are described as follows.

STEP 1. Perform the image preprocessing. Usually geometric distortion exists in the image taken; therefore, it is necessary to correct each image by applying the geometric correction method [9].

The Sobel edge operator is used to detect edge points in the corrected image, and the fast parallel thinning algorithm proposed by Zhang and Suen [10] is used to thin the edge points found by the Sobel operator.

STEP 2. Locate L_a and L_b in the thinned image. The Hough transform [11] is applied for line detection.

STEP 3. Use least-square-error line fitting [12] to improve the accuracy of L_a and L_b. The set of line points (denoted by $\{(u_i, v_i) \mid i = 1, 2, \ldots, n\}$ to be fitted is composed of the points near the line obtained in Step 2. By the this fitting process, more accurate equations of L_a and L_b can be obtained. We then choose two points A' and B' from L_a and two points C' and D' from L_b to find $\hat{N}_1$ and $\hat{N}_2$.

STEP 4. Detect the upper and the lower curves (i.e., the projections of the upper and the lower edges of the generalized cylindrical portion) in the thinned image.

STEP 5. Adopt a two-stage detection procedure [8] to find two exact point-pairs. In the first stage, for each pixel P_i on the upper curve, the point-pair estimation algorithm is employed to find the corresponding pixel P_j on the lower curve with the smallest Δ and $[P_i, P_j]$ is recorded as a candidate point-pair. In the second stage, two local curves (called upper local curve α and lower local curve β) are derived by using the least-square error curve fitting to fit two sets of pixels on the upper and lower curves, respectively. α is approximated using pixel P_i and its neighbor pixels P_{i+k}, $k = -3 \sim 3$, and β is approximated using pixel P_j and its neighbor pixels P_{j+k}, $k = -3 \sim 3$. Assume the arc length of the curves α and β are L_α and L_β, respectively, and the sampling rate is $\Delta s_1 = L_\alpha / N'$ and $\Delta s_2 = L_\beta / N'$, where N' is an integer (in our case, $N' = 30$). The point-pair estimation algorithm was applied to those points $P_i + m'\Delta s_1$ on the curve α and $P_j + m'\Delta s_2$ on the curve β, where $m' = -15 \sim 15$. The point P_{ii} on the curve α and P_{jj} on the curve β with the minimum Δ were chosen as a desired point-pair $[P_{ii}, P_{jj}]$. The other point-pair can be obtained by a similar procedure.

The above image procedures is performed under the assumption that there is only one cylindrical object in the image. If there are multiple cylindrical objects in the image and one is partially occluded by another, the above procedures should be replaced by more complicated methods [8]. The details of those methods are omitted here.

Once the required features are extracted, the robot location relative to the object can be calculated from the formulas derived in Section 2.

4. EXPERIMENTAL RESULTS AND DISCUSSIONS

Images were processed by a Series 150/151 modular image processor connected to a 80386 PC/AT and programs were written in C language. In the experiments, each tested object was placed at several different locations and two pictures were taken for each object at each location. The real robot location $(0, Y_c, Z_c)$ was manually measured as the reference for checking location results. The two point-pairs with small values of the minimized discriminant Δ (denoted by $\Delta_{\min}$) described in Section 2 were chosen for constructing the rectangle-shaped standard mark. By using the two-stage point-pair estimation algorithm, two point-pairs with the smallest $\Delta_{\min}$ and separated to some extent were then used to construct the standard mark. Table 1 shows the experimental result of robot location for each object at each location. In the table, the unit is cm and the last column shows the error percentage for each object compared with the real reference position and according to the results, all the average error percentages can be seen to be less than 5%, which shows that the approach is feasible for practical applications.

Correct focusing is important because an image out of focus will contain blurred edges which are fatal to the robot location determination. This problem can be solved by using autofocusing cameras. In real applications, errors are introduced into the computation owing to the limitation of sensor resolution and noises. The effect resulting from the finite resolution is that a small area

Table 1. The experimental results of robot location.

Test Object	Measured location (cm)	Computed location (cm)	Error percentage (%)
Triangular cylinder	(0, 50.0, 32.0)	(0.3, 52.1, 31.1)	(3.0, 4.2, 2.8)
	(0, 55.0, 30.0)	(0.2, 56.4, 31.1)	(2.0, 2.5, 3.6)
Rectangular cylinder	(0, 67.0, 25.0)	(0.3, 67.9, 24.4)	(3.0, 1.3, 2.4)
	(0, 61.0, 32.0)	(0.4, 62.2, 31.5)	(4.0, 1.9, 1.5)
Elliptic cylinder	(0, 80.0, 40.0)	(0.2, 81.6, 41.5)	(2.0, 2.0, 3.7)
	(0, 77.0, 35.0)	(0.3, 75.2, 36.1)	(3.0, 2.3, 3.1)
Circular cylinder	(0, 70.0, 28.0)	(0.2, 72.1, 28.5)	(2.0, 3.0, 1.8)
	(0, 80.0, 38.0)	(0.3, 83.1, 37.2)	(3.0, 3.8, 2.1)
Hexangular cylinder	(0, 85.0, 26.0)	(0.5, 82.7, 26.4)	(5.0, 2.7, 1.5)
	(0, 78.0, 26.0)	(0.3, 76.4, 26.7)	(3.0, 2.0, 2.6)

in 3D space may be projected onto the same point in the image; the farther the object from the camera is, the more serious this effect will be. The error caused by the small deviation of the image plane center is negligible [13]. In addition, the glare reflected from the material from which the label is made can also introduce errors. This problem can be alleviated by using materials which attenuate glare.

5. CONCLUSIONS AND SUGGESTIONS

A feasible approach to the problem of determining the robot location relative to a generalized cylindrical object by a single image is proposed in this paper. The principle idea of the approach is to use a rectangle-shaped standard mark for performing monocular image analysis. The standard mark is constructed by two point-pairs found by the two-stages point-pair estimation. The derivation involved in the proposed approach consists of mere 3D vector analysis and simple algebraic computation. Owing to the linear derivation, this approach can achieve high speed requirement. Moreover, since the camera focal length can be derived, there is no need to perform the calibration of the camera focal length. This is especially useful in the sensor systems where autofocusing cameras are used. In addition, no prerequisite is imposed on the image-grabbing process. Experimental results show that the location determination time is about 1.2 sec in a 33 MHz IBM compatible PC/AT computer system and the location error is less than 5% in average.

Much time is consumed in image preprocessing, which includes geometric correction, edge detection and thinning. If the execution time is concerned, the above three processes must be implemented in hardware. To reach higher accuracy, we may use imaging devices with higher resolution and perform more accurate geometric correction.

Further research includes generalizing the method to objects of other shapes using standard marks of other forms.

APPENDIX

In Figure 4, planes S_1 and S_2 are tangent to the generalized cylindrical object and pass through the camera lens center O'. Note that S_1 is the plane that contains O', A' and B', and S_2 is the plane that contains O', D' and C'. Points A' and B' (C' and D') are the projections of points A and B (C and D), which lie on the intersection line of S_1 (S_2) and the generalized cylindrical object. Let $\hat{N}_1$ and $\hat{N}_2$ denote the normal vectors of S_1 and S_2, respectively, and (u_a, v_a), (u_a, v_b),

(u_c, v_c) and (u_d, v_d) be the image coordinates of A', B', C' and D', respectively. Then $\hat{N}_1$ and $\hat{N}_2$ can be obtained from

$$\hat{N}_1 = \overrightarrow{O'B'} \times \overrightarrow{O'A'} = \begin{vmatrix} \hat{x}' & \hat{y}' & \hat{z}' \\ u_b & f & v_b \\ u_a & f & v_a \end{vmatrix} = (a_1 f, b_1, c_1 f),$$

and

$$\hat{N}_2 = \overrightarrow{O'D'} \times \overrightarrow{O'C'} = \begin{vmatrix} \hat{x}' & \hat{y}' & \hat{z}' \\ u_d & f & v_d \\ u_c & f & v_c \end{vmatrix} = (a_2 f, b_2, c_2 f),$$

where $a_1 = v_a - v_b$, $b = u_a v_b - u_b v_a$, $c_1 = u_b - u_a$, $a_2 = v_c - v_d$, $b_2 = u_c v_d - u_d v_c$ and $c_2 = u_d - u_c$. Figure 4 shows clearly that the direction of the object axis (denoted by $\hat{N}$) is given by

$$\hat{N} = \hat{N}_1 \times \hat{N}_2 = \begin{vmatrix} \hat{x}' & \hat{y}' & \hat{z}' \\ a_1 f & b_1 & c_1 f \\ a_2 f & b_2 & c_2 f \end{vmatrix} = f(a_3, b_3 f, c_3),$$

where $a_3 = b_1 c_2 - b_2 c_1$, $b_3 = a_2 c_1 - a_1 c_2$ and $c_3 = a_1 b_2 - a_2 b_1$. The multiplier f before the vector $(a_3, b_3 f, c_3)$ can be neglected.

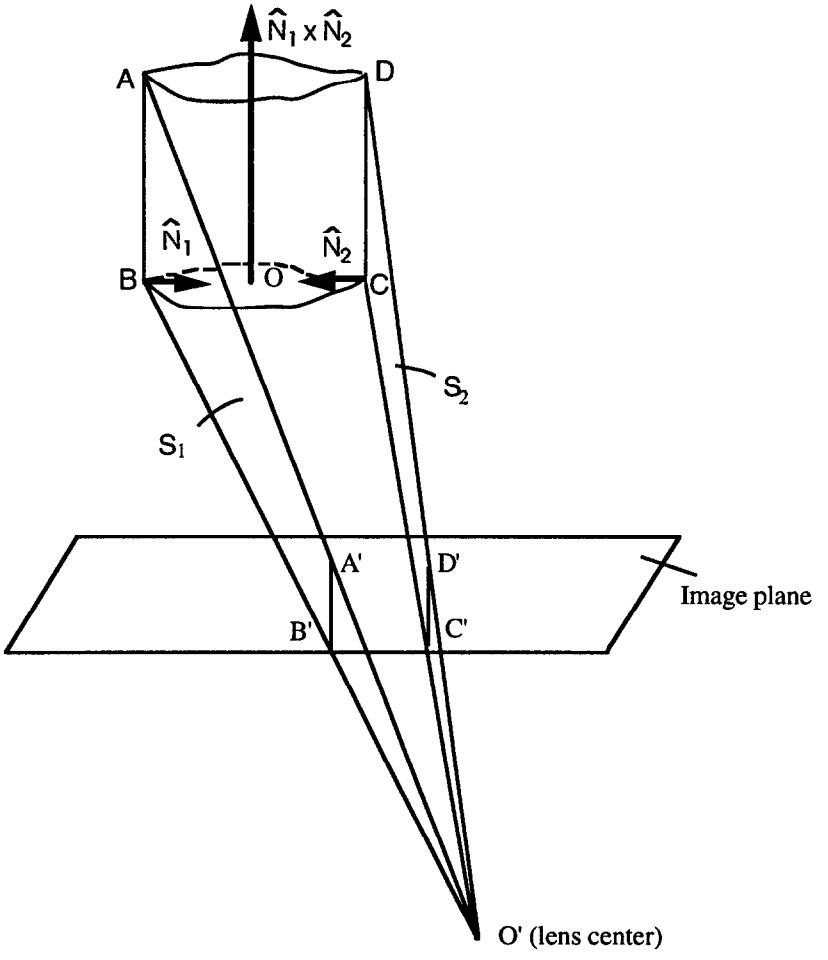

Figure 4. Two tangent planes S_1 and S_2 with normals $\hat{N}_1$ and $\hat{N}_2$, respectively.

REFERENCES

1. I. Fukui, TV image processing to determine the position of robot vehicle, *Pattern Recognition* **14**, 101–109 (1981).
2. J.W. Courtney, M.J. Magee and J.K. Aggarwal, Robot guidance using computer vision, *Pattern Recognition* **17**, 585–592 (1984).
3. C.N. Shyi, A study of mark-based robot vision system, Doctoral Dissertation, National Cheng Keng University, (1989).
4. K.C. Hung, C.N. Shyi, J.Y. Lee and T.C. Lee, Robot location determination in a complex environment by multiple marks, *Pattern Recognition* **21**, 567–580 (1988).
5. H.L. Chou and W.H. Tsai, A new approach to robot location by house corners, *Pattern Recognition* **19**, 439–451 (1986).
6. R.M. Haralick and Y.H. Chu, Solving camera parameters from the perspective projection of a parameterized curve, *Pattern Recognition* **17**, 637–645 (1984).
7. S.Y. Chen and W.H. Tsai, Robot location using surface patches of curved objects, *Int. J. of Robotics and Automation* **4**, 123–134 (1990).
8. J.-D. Lee, A study on recognition and localization of common curved objects, Doctoral Dissertation, National Cheng Kung University, (1992).
9. A. Rosenfeld and A.C. Kak, *Digital Picture Processing*, Vol. 2, Academic Press, New York, (1982).
10. T.Y. Zhang and C.Y. Suen, A fast parallel algorithm for thinning digital patterns, *Commun. ACM* **27** (3), 236–244 (1984).
11. D.H. Ballard, Generalizing the Hough transform to detect arbitrary shapes, *Pattern Recognition* **13**, 111–122 (1983).
12. M.L. James, *Applied Numerical Methods for Digital Computation*, (3rd Edition), The Maple Press, New York, (1985).
13. R.Y. Tsai, A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shell TV cameras and lens, *IEEE J. Robotics and Automation* **3**, 323–344 (1987).